

APPLICATION
FOR
UNITED STATES LETTERS PATENT

TITLE: METHOD FOR MUTAGENESIS
APPLICANT: ATSUSHI MIYAWAKI AND ASAKO SAWANO

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EF353814376US

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

August 2, 2001

Date of Deposit

Signature

Francisco Robles

Typed or Printed Name of Person Signing Certificate

METHOD FOR MUTAGENESIS

FIELD OF THE INVENTION

The present invention relates to an entirely new method for mutagenesis, which can be applied for introducing certain mutations at certain sites of a nucleotide sequence, or for introducing random mutations at certain sites of a nucleotide sequence.

BACKGROUND OF THE INVENTION

The method for site-directed mutagenesis is a powerful tool for producing mutants to assess the importance of specific amino acid residues in a protein's structure and function. As such method for site-directed mutagenesis, a variety of protocols have been established, including those using the polymerase chain reaction (PCR).

Among the PCR-based protocols, the Quik Change™ Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) (referred to as "Quik Change method", hereinafter) is widely used. According to the protocol of said Quik Change method, a DNA containing mutations can be obtained through the following three steps of:

- 1) carrying out the PCR process by use of a pair of complementary primers designed to carry a mutation at a desired site;
- 2) digesting only the template DNAs (that carry no mutations) in the reaction solution; and
- 3) performing transformation by use of the DNA fragments which are not digested in the above step of digestion.

In particular, said Quik Change method utilizes, in the step of carrying out the PCR process, a supercoiled, double-stranded (ds) DNA plasmid as a template DNA, and a high-fidelity Pfu Turbo™ DNA polymerase as an enzyme. During the step of the PCR process, the pair of complementary primers is annealed to the dsDNA, followed by an elongation reaction so as to amplify DNA fragments that carry mutations and are nicked between the elongated terminus and the 5'-terminus of the primers.

Subsequently, in the Quik Change method, DpnI endonuclease is caused to act in the reaction solution to specifically cut fully- or hemi-methylated 5'-GATC-3' sequences. The *in vitro* synthesized and nicked plasmid DNA including the desired mutation is then transformed into *E. coli*. As a result, the plasmid DNA that is nicked will be repaired in *E. coli* and become circular.

Thus, according to the Quik Change method, it is possible to synthesize a DNA including a desired mutation at a desired site. However, the Quik Change method suffers from the following disadvantages:

- 1) multiple mutations can not be introduced at different relatively separated positions in a nucleotide sequence at the same time, i.e., in this method, it is possible to introduce multiple mutations only in a pair of complementary primers, whereas it is impossible to introduce multiple mutations at remote positions;
- 2) when introducing a mutation, a pair of synthetic complementary oligonucleotides primers is required, that is, this method requires two primers for introducing a mutation at one site; and
- 3) since it requires a pair of synthetic complementary oligonucleotides primers, it does not allow random mutagenesis at desired sites using degenerative primers.

The Quik Change method, due to the above problem 1), requires that the aforementioned 3 steps be repeated the number of times corresponding to the number of mutagenesis, when introducing mutations at multiple sites. Consequently, the Quik Change method takes a long time to introduce mutations at multiple sites, leading to an inconvenience of poor efficiency. Additionally, the Quik Change method, owing to the above problem 2), involves high cost in synthesizing or purchasing the pair of synthetic oligonucleotides primers. Moreover, the Quik Change method, due to the above problem 3), can not be used in experimental systems where certain amino acid residues in the analyte protein are mutated into a variety of other amino acids, thus limiting the experimental systems to which it is applicable.

SUMMARY OF THE INVENTION

Thus, in view of the problems of prior art mentioned above, an object of the present invention is to provide an entirely new method for mutagenesis, which method being simple, speedy, economical, and widely-applicable.

The method for mutagenesis according to the present invention, having achieved said object, comprises the following steps of:

a DNA synthesis in which one or more primers which have a nucleotide sequence containing at least one mutation and a phosphorylated 5'-terminus are annealed to a template DNA and then subjected to an elongation reaction using a thermostable high-fidelity DNA polymerase, after which the phosphorylated 5'-terminus and the elongated terminus are ligated by means of a thermostable DNA ligase to synthesize a circular DNA containing said primers;

a digestion in which said step of DNA synthesis is repeated several times to amplify the DNA containing said primers, after which at least the DNAs other than the amplified circular DNA are then digested into several fragments; and

a double-stranded DNA synthesis in which, with the several fragments obtained in the above step of digestion as megaprimers, said megaprimers are annealed to said circular DNA synthesized in the above step of DNA synthesis, followed by an elongation reaction performed using said thermostable high-fidelity DNA polymerase.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an overview illustrating the protocol of a method for mutagenesis according to the present invention.

Figure 2 shows the fluorescence excitation (A) and emission (B) spectra of EGFP, ECFP, EYFP and EGFP-Y66WT203Y.

Figure 3 shows the excitation (broken line) and emission (solid line) spectra of ECFP,

ECGFP and EGFP.

Figure 4 shows the pH-dependency of the excitation and emission spectra of EGFP, ECFP and ECGFP respectively, where (A): the absorbance spectrum of EGFP, (B): the emission spectrum of EGFP, (C): the absorbance spectrum of ECFP, (D): the emission spectrum of ECFP, (E): the absorbance spectrum of ECGFP, and (F): the emission spectrum of ECGFP.

Figure 5 shows the pH-dependency of the absorbance and emission peaks of EGFP, ECFP and ECGFP respectively. In the figure, (A) illustrates the relationship between the absorbance peak and pH and (B) illustrates the relationship between the fluorescence peak and pH.

EMBODIMENT OF THE PRESENT INVENTION

The method of mutagenesis according to the present invention will be described below in detail with reference to the attached drawings. Following examples will illustrate a method for introducing mutations at certain sites to a gene encoding the enhanced green fluorescent protein (referred to as EGFP hereinafter) which is made by altering the green fluorescent protein (GFP) produced by *Aequorea*. The method for mutagenesis according to the present invention is a method that is applicable not only for introducing mutations to said gene but to any genes.

The gene encoding said EGFP (SEQ ID NO.1) used herein is a gene which is changed in codon usage as compared with the gene encoding GFP (SEQ ID NO.3), and enhanced in expression within a host by adding a codon encoding valine between the first methionine and the second serine. EGFP (whose amino acid sequence is shown in SEQ ID NO.2) is made by replacing phenylalanine at position 64 in GFP (whose amino acid sequence is shown in SEQ ID NO.4) by leucine (at position 65 in EGFP) and serine at position 65 by threonine (at position 66 in EGFP). In the present specification, for indicating the position of an amino acid, all of the numerical values should be

construed as those numbered on the basis of the amino acid sequence of GFP. Thus, for example “tyrosine at position 66 in EGFP” means the amino acid at position 67 in SEQ ID NO.2. In the present method, as shown in Fig. 1, mutations are introduced in a gene encoding EGFP through the first step of carrying out “synthesis of a mutated strand”, the second step of carrying out “digestion with of DpnI” and the third step of carrying out “synthesis of a complementary strand by use of megaprimers”.

(1) “Synthesis of a mutated strand”

First of all, primers for mutagenesis are prepared to introduce mutations at certain sites in the nucleotide sequence encoding known EGFP (SEQ ID NO.1). As the present method allows multiple mutations to be introduced at a time, the primers for mutagenesis are prepared in a number corresponding to the number of sites at which mutations are introduced.

These primers for mutagenesis are designed on the basis of SEQ ID NO.1 and at the same time are designed as those replacing certain bases in SEQ ID NO.1. The primers for mutagenesis are not limited to those replacing a single base and may be those replacing several bases. In this case, the primers for mutagenesis are designed such that they can be annealed to a plasmid containing a gene encoding EGFP.

The 5'-termini of these primers for mutagenesis have been phosphorylated using, for example, T4 polynucleotide kinase. Here, the enzyme that phosphorylates the 5'-termini of these primers for mutagenesis (after-mentioned Pri-1 or Pri-2), being not limited to T4 polynucleotide kinase, may include any enzymes capable of phosphorylating the 5'-termini of the primers.

Specifically, as a primer for mutagenesis, the following may be given by way of example: Pri-1 (5'-GCACTGCACGCCCCAGGTCAGGGTGGT-3' : SEQ ID NO.5) for mutating tyrosine (Y) at position 66 into triptophan (W) and Pri-2 (5'-GGGCGGACTGGTAGCTCAGGTAGTGG-3' : SEQ ID NO.6) for mutating threonine

(T) at position 203 into tyrosine (Y).

In addition, a plasmid containing genes encoding EGFP is prepared. As such plasmid, for example, pRSET_B (manufactured by Invitrogen Co.) having a T7 expression system may be used. At this time, a well-known approach can be used to construct pRSET_B (pRSET_B/EGFP) containing genes encoding EGFP. The plasmids are normally methylated and the pRSET_B/EGFP used here is also methylated.

When synthesizing the mutated strand, a PCR process is carried out that uses these primers for mutagenesis (Pri-1 and Pri-2) and the relevant plasmid (pRSET_B/EGFP). In this process, a thermostable high-fidelity DNA polymerase is used in the elongation reaction from the primers for mutagenesis using a plasmid as a template. This thermostable high-fidelity DNA polymerase is that having about 10 times higher fidelity than other thermostable DNA polymerases such as TaqDNA polymerase, and capable of synthesizing a complementary strand for the template DNA substantially without mistake, for it has practically no activity to add excessively on the order of one base to the 3'-terminus of elongated strand. As a thermostable high-fidelity DNA polymerase, Pfu DNA polymerase (manufactured by Stratagene Co., La Jolla, CA) may be used. The thermostable high-fidelity DNA polymerase is not limited to Pfu DNA polymerase and may include any thermostable high-fidelity DNA polymerase as long as it has a high-fidelity and practically no activity to add excessively on the order of one base to the 3'-terminus of elongated strand.

In the PCR process, after the elongation reaction using the thermostable high-fidelity DNA polymerase, the 5'-terminus of a primer for mutagenesis and the 3'-terminus of an elongated DNA strand are ligated with a thermostable DNA ligase. As the thermostable DNA ligase, while not being specially limited, for example Taq ligase (New England Biolabs) may be used. In the present method, phosphorylation of the 5'-termini of the primers for mutagenesis allows the ligation as described above to be achieved.

The thermal cycler in said PCR process was programmed as follows: pre-incubation at 65°C for 5 min allowing Taq DNA ligase to repair any nicks in the template DNA; initial denaturation at 95°C for 2 min; 18 cycles, each cycle consisting of DNA denaturation at 95°C for 30 sec, annealing reaction at 55°C for 30 sec and elongation and ligation reaction at 65°C for 7 min; and finally, post-incubation at 75°C for 7 min. In this program, the time for the elongation and ligation reaction at 65°C may be set longer to complete the ligation by means of Taq DNA ligase.

A program such as that described above enables to amplify only the single-stranded plasmid DNA containing a mutated EGFP gene made by introducing mutations that the primers for mutagenesis carry into a gene encoding EGFP. Specifically, the base sequence of a mutated EGFP gene synthesized using Pri-1 and Pri-2 as the primers for mutagenesis is shown in SEQ ID NO.8.

The final reaction solution of said PCR process comprises a single-stranded plasmid DNA containing the mutated EGFP gene, a single-stranded DNA derived from the plasmid being a template, a double-stranded plasmid being a template, and a double-stranded DNA made by annealing the single-stranded plasmid DNA containing the mutated EGFP gene and the single-stranded DNA derived from the plasmid being a template.

(2) "Digestion with DpnI"

In this step, at least an enzyme that digests DNA other than the single-stranded DNA containing the mutated EGFP gene is added to the reaction solution obtained as described above, to digest and remove only certain DNAs. As way of example, enzymes that can be used in this step include DpnI that recognizes and cuts certain nucleotide sequences (5'-G^mATC-3').

When DpnI is added to the reaction solution, it will selectively cut the methylated

plasmid (pRSET_B/EGFP) remaining in the solution, and the hemi-methylated double-stranded DNA, made by annealing the single-stranded plasmid DNA containing the mutated EGFP gene and the single-stranded DNA derived from the plasmid being a template, remaining in the solution.

As a result, in the reaction solution, there exist as circular DNAs only the single-stranded DNA containing the mutated EGFP gene and a trace of the single-stranded DNA derived of plasmid being a template.

(3) "Synthesis of complementary strand by use of megaprimers"

In this step, the DNA fragments cut in the above step (2) are used as megaprimers and the single-stranded plasmid DNA containing the mutated EGFP gene is used as a template to synthesize a double-stranded DNA. Namely, in this reaction, the DNA fragments cut by DpnI are denatured to make single-stranded DNA fragments, while the single-stranded DNA fragments are annealed to the single-stranded plasmid DNA containing the mutated EGFP, and then an elongation reaction is performed using said single-stranded DNA fragments as megaprimers. This results in the synthesis of a nucleotide sequence complementary to the single-stranded plasmid DNA containing the mutated EGFP gene, so as to construct a double-stranded plasmid DNA containing EGFP genes in which mutations are introduced at desired sites.

Specifically, the megaprimers are denatured at 95 °C for 30 sec to become single-stranded, after which 2 cycles are carried out, each cycle consisting of thermal denaturation at 95°C for 30 sec, annealing at 55°C for 1 min and elongation reaction at 70°C for 7 min, to synthesize a double-stranded DNA.

The DNA fragments obtained by treatment with DpnI include those containing wild-type bases present before the introduction of mutations. When using, as megaprimers, DNA fragments in which the sites to introduce mutations are wild-type nucleotide sequences, mutations will not be introduced into the synthesized

complementary strand. Therefore, in this step, a large amount of an auxiliary primer is added in order to synthesize the double-stranded DNA in such a state that annealing of the relevant DNA fragment is prevented.

As an auxiliary primer, T7 primer can be used, which is complementary to the sequence present in the vicinity of the cloning site of the EGFP gene. Addition of T7 primer to the reaction solution allows the T7 primer to anneal preferentially, thus preventing the annealing of the DNA fragments in which the sites to introduce mutations are wild-type nucleotide sequences.

In this step, it is also possible to utilize Pfu DNA polymerase, Taq DNA ligase, dNTPs and NAD carried over from the PCR in the above step (1). Namely, in the present method, since the reaction can be carried out in the same reaction solution as used in the above steps (1) and (2) and continuously after these steps, the desired double-stranded plasmid DNA can be synthesized without going through a complex process such as purification of DNA fragments.

As seen from the above, the steps (1), (2) and (3) together make it possible to construct the double-stranded plasmid containing the mutated EGFP gene made by introducing mutations at certain sites in the EGFP gene. Further, certain cells are transformed using the reaction solution containing the relevant double-stranded plasmid to obtain transformed cells comprising plasmid containing the mutated EGFP gene.

The cells that can be used herein include any of the following cells such as bacteria like *E. coli* and *Bacillus subtilis*, yeast cells, animal cells, insect cells and plant cells etc. The cells may be appropriately selected depending on the origin of the gene in which mutation is introduced or on the plasmid which is used. Also, in said transformation, any of the generally used approaches may be applied, for example, that using competent cells, that applying an electric pulse, or the like.

The mutated gene can be expressed by culturing the resulting transformed cells. For example, in the case of using pRSETb/EGFP as a template DNA, *E. coli* JM109(DE3) is transformed and isopropyl β -D-thiogalactoside (IPTG) is added in the medium so as to induce expression of mutated EGFP.

Mutated protein, which is expressed in the transformed cells, can be purified by subjecting it to any of the protein purification processes that are generally used. For purifying the mutated protein, one may use, for example, affinity chromatography, gel filtration or the like.

However, in the above step (1) "synthesis of a mutated strand", the method for mutagenesis according to the present invention is not limited to those using multiple primers for mutagenesis, but also include those using degenerative primers to introduce random mutations into certain nucleotide sequences. The following example will illustrate a method wherein degenerative primers are used to mutate randomly amino acids at certain sites in ECFP (enhanced cyan fluorescent protein obtained by altering cyan fluorescent protein which is produced by *Aequorea*).

In this case, for example, as a degenerative primer, a degenerative primer Pri-3 (5'-GCGGACTGNNNGCTCAGGTAG-3': SEQ ID NO.7) is synthesized for mutating randomly threonine (T) at position 203 in the EGFP amino acid sequence. Here, "N" in the nucleotide sequence of Pri-3 (SEQ ID NO.7) means A, G, C or T, and can be synthesized by use of a nucleoside phosphoramidite containing A, G, C and T at equimolar concentrations.

In the above step (1) "synthesis of a mutated strand", Pri-3 can be used to construct a variety of mutated ECFP genes in which threonine (T) at position 203 in EGFP is randomly mutated.

EXAMPLES

The present invention will be described in more detail in the following examples, which, however, do not limit in any way the technical scope of the invention.

Example 1

1) Preparation of primers for mutagenesis and template plasmid DNA

In this example, tyrosine (Y) at position 66 is mutated into triptophan (W) and threonine (T) at position 203 into tyrosine (Y). As the primers for mutagenesis, Pri-1 (5'-GCACTGCACGCCCCAGGTCAGGGTGGT-3' : SEQ ID NO.5) for mutating tyrosine (Y) at position 66 into triptophan (W) and Pri-2 (5'-GGGCGGACTGGTAGCTCAGGTAGTGG-3' : SEQ ID NO.6) for mutating threonine (T) at position 203 into tyrosine (Y) were designed and synthesized. The 5'-termini of said Pri-1 and Pri-2 were phosphorylated respectively with T4 polynucleotide kinase.

To prepare a template plasmid DNA, the EGFP gene was cloned into a plasmid vector pRSET_B (manufactured by Invitrogen Co.). The template plasmid DNA used hereafter is referred to as "pRSET_B/EGFP".

2) Synthesis of a mutated strand

50 μ l of a reaction solution was prepared to carry out PCR:

pRSET _B /EGFP	50 ng
Pri-1	14 pmol
Pri-2	14 pmol
dNTPs	10 nmol
Pfu DNA polymerase	2.5 U
Taq DNA ligase	20 U

The Pfu DNA polymerase is that manufactured by Stratagene Co., and dissolved in a 0.5 x Pfu polymerase reaction buffer. In addition, the Taq DNA ligase is that manufactured by New England Biolabs Co., and dissolved in a 0.5 x Taq DNA ligase

buffer containing 50 nmol of NAD.

The thermal cycler was programmed as follows: pre-incubation at 65°C for 5 min allowing Taq DNA ligase to repair any nicks in the template DNA; initial denaturation at 95°C for 2 min; 18 cycles, each cycle consisting of DNA denaturation at 95°C for 30 sec, annealing reaction at 55°C for 30 sec and elongation and ligation reaction at 65°C for 7 min; and finally, post-incubation at 75°C for 7 min.

3) Digestion with DpnI

DpnI was used under the following conditions to digest selectively the methylated or hemi-methylated plasmid DNAs. Namely, 1 μ l (20 U) of DpnI (manufactured by New England BioLabs Co.) was added to 50 μ l of the reaction solution after the above PCR process. This solution was then incubated at 37°C for 1 hour.

DpnI is an endonuclease which recognizes (5'-G^mATC-3') to cut double-stranded DNAs. This DpnI completely cuts the plasmid pRSET_B/EGFP of 3672 bp to produce 23 DNA fragments of 6, 8, 11, 12, 17, 18, 24, 31, 36, 38, 46, 47, 75, 78, 105, 133, 148, 160, 258, 341, 506, 740, and 789 base pairs.

4) Synthesis of complementary strand by use of megaprimers

51 μ l of the reaction solution after digestion by DpnI was used to synthesize a complementary strand for the mutation strand synthesized in the above step 2). That is, 51 μ l of the reaction solution was denatured at 95°C for 30 sec, after which 2 cycles at 95°C for 30 sec, 55°C for 1 min, and 70°C for 7 min were carried out.

At this time, some of the 23 DNA fragments obtained in the above step 3) served as megaprimers.

5) Transformation and expression of mutated EGFP

2 μ l of the sample after the reaction in the above step 4) was used to transform *E. coli*

competent cell (JM109(DE3)) according to Ca^{2+} coprecipitation. Then, the transformed *E. coli* was cultured at 37°C in a LB solid medium containing 100 $\mu\text{g/ml}$ ampicillin. Thereafter, grown single colonies were picked up and transferred to 2 ml of a LB liquid media containing 100 $\mu\text{g/ml}$ ampicillin, which was cultured at 37°C overnight to grow.

The mutated EGFP was expressed using the T7 expression system (pRSETb/JM109(DE3)). Namely, a 100-fold dilution of the culture was prepared and grown at room temperature until said culture reached a density of approximately $\text{OD}_{600}=0.5$, and then expression of the mutated EGFP was induced by isopropyl β -D-thiogalactoside (IPTG) at a final concentration of 0.2mM, followed by further culturing for 18 hours. The resultant cells were lysed by means of a French press, and the polyhistidine tagged proteins were purified from the cleared lysates using a nickel-chelate column (manufactured by Qiagen Co.). In addition, the protein samples in the eluates (100 mM imidazole, 50 mM Tris/Cl pH 7.5, 300 mM NaCl) were concentrated by Centricon 10 (Amicon) and were further purified by gel filtration.

6) Results

1. It was confirmed that mutations were introduced simultaneously at two sites in the resultant proteins by measuring the excitation wavelength and emission length of said proteins, and comparing them with those of existing proteins (such as wild-type EGFP).

In Fig. 2 (A), Curve A represents the excitation wavelength measured and normalized on the resultant proteins. In Fig. 2 (B), Curve A represents the emission wavelength measured and normalized on the resultant proteins. Further, Figs. 2 (A) and (B) also show, for the purpose of comparison, the excitation and emission wavelengths on the wild-type EGFP-Y66W made by mutating tyrosine at position 66 into triptophan and the wild-type EGFP-T203Y made by mutating threonine at position 203 into tyrosine. In Figs. 2 (A) and (B), Curves B, C and D represent the wild-type EGFP, EGFP-Y66W and EGFP-T203Y, respectively.

As shown in Figs. 2 (A) and (B), it is found that the proteins in which mutations were introduced by the aforementioned method have an excitation wavelength peak at 458 nm and an emission wavelength peak at 504 nm. On the other hand, the wild-type EGFP has an excitation wavelength peak at 488 nm and an emission wavelength peak at 509 nm. EGFP-Y66W, producing a new chromophore containing indole by replacing tyrosine at position 66 by triptophan, has an excitation wavelength peak at 450 nm and an emission wavelength peak at 494 nm. EGFP-T203Y, producing a $\pi \cdot \pi$ interaction with tyrosine at position 66 by replacing threonine at position 203 by tyrosine, has an excitation wavelength peak at 514 nm and an emission wavelength peak at 524 nm.

The resultant proteins, being made by mutating tyrosine at position 66 into triptophan and mutating threonine at position 203 into tyrosine (EGFP-Y66W/ T203Y), produce said $\pi \cdot \pi$ interaction and said chromophore containing indole, thus present a cyan-green luminescent color.

2. On the other hand, the mutation introduction rate is calculated by measuring the fluorescence of the colonies grown on the LB solid media containing ampicillin. Spectral properties of the relevant colonies were measured using a fluorescence image analysis system, which was adapted on the basis of the report by Baird et al. (Baird, G.S., Zacharias, D.A. and Tsien, R.Y. (1999) Proc. Natl. Acad. Sci. USA, 96, 11241-11246). In said image analysis system, the colonies grown on the LB plate (diameter of 9 cm) were excited by means of a 150 W xenon lamp. The excitation wavelength can be selected by a diffraction grating monochrometer (CT-10, manufactured by JASCO Co.). Specifically, the excitation wavelengths selected by said diffraction grating monochrometer are set to be 450 nm and 490 nm (indicated by the straight lines a and b in Fig.2 (A)).

Moreover, in this system, a pair of optical fibers (1 m) was used in order to illuminate

the entire LB plate evenly. The fluorescence images from the colonies were collected through an interference band-pass filter (535DF25, 1 inch in diameter, manufactured by Omega Co.) into a lens (AF NIKKOR, manufactured by Nikon Co.). The images were captured by a SenSys cooled CCD camera (manufactured by Photometrics Co.) and analyzed using MetaMorph 3.0 software (manufactured by Universal Imaging Co., West Chester, PA). Specifically, the ratio image was obtained by dividing the 490 nm image by the 450 nm image. As expected from their excitation spectra shown in Fig.2 (A), the ratio values were as follows: EGFP-T203Y > wild-type EGFP > EGFP-Y66W/T203Y > EGFP-Y66W.

At this time, when comparing with the interference band-pass filter, use of the diffraction grating monochromator permitted us to change and select excitation wavelengths of very narrow width. Thus, in this system, the excitation using said monochromator enabled to definitely distinguish the wild-type EGFP, EGFP-T203Y, EGFP-Y66W and EGFP-Y66W/T203Y respectively, each of which presented close excitation spectra.

As a result, out of 42 colonies grown on the LB plate containing ampicillin, 32 of them showed cyan-green fluorescence (EGFP-Y66W/T203Y), and the efficiency of simultaneous mutagenesis at 2 sites was calculated to be 76 %. The rest of the colonies were as follows: 4 showed yellow fluorescence (EGFP-T203Y); 4 showed cyan fluorescence (EGFP-Y66W); and 2 showed green fluorescence (wild-type EGFP). There were no non-fluorescent colonies.

Further, the experiment described above was repeated 4 times so as to calculate the rate of colonies showing cyan-green fluorescence, i.e., the efficiency of simultaneous mutagenesis at 2 sites. As a result, (number of colonies showing cyan-green fluorescence) / (total number of colonies) was 21/30, 42/50, 50/72 and 32/40, respectively and the efficiency of simultaneous mutagenesis at 2 sites was found to be $75.8 \% \pm 4.2 \% (n=5)$.

Also, by DNA sequencing of the plasmid purified from the colonies showing cyan-green fluorescence, it was found that there were no mutations introduced other than at the sites of Pri-1 and Pri-2.

Moreover, out of 23 DNA fragments serving as megaprimers in the step 4), the sense strands of 506 bp and 148 bp contain a region encoding the amino acid residue 66 and a region encoding the amino acid residue 203, respectively. For this reason, it is preferable that these DNA fragments of 506 bp and 148 bp do not serve as megaprimers. In the present example, addition of a common sense primer (T7 primer 14 pmol) that anneals outside the cloning site after DpnI digestion, enabling prevention of the annealing of the DNA fragments of 506 bp and 148 bp, can increase the efficiency of the mutagenesis to over 90 %. Specifically, out of 200 colonies, 196 showed cyan-green fluorescence.

Example 2

In this example, a degenerative primer was used to mutate randomly threonine (T) at position 203 in the ECFP amino acid sequence. As the primer for mutagenesis, a degenerative primer Pri-3 (5'-GCGGACTGNNNGCTCAGGTAG-3': SEQ ID NO.7) was synthesized to mutate randomly threonine (T) at position 203, wherein "N" in the nucleotide sequence of Pri-3 (SEQ ID NO.7) means A, G, C or T. Then, the 5'-termini of these Pri-3 were respectively phosphorylated with T4 polynucleotidase kinase.

As the template plasmid DNA, an ECFP gene was cloned into a plasmid vector pRSET_B (manufactured by Invitrogen). Hereafter the template plasmid DNA used is "pRSET_B/ECFP". This ECFP gene has additional substitutions (N146I, M153T, V163A and N164H) to facilitate folding and maturation of the proteins, with tyrosine at position 66 being replaced by tryptophan (Y66W).

For Example 2, the same experiment as in Example 1 was carried out, except that

Pri-3 was used as the degenerative primer and pRSET_B /ECFP was used as the template DNA.

1) Results

In this example, after the transformation carried out as described in Example 1, the transformed *E. coli* cell was grown on a LB plate containing ampicillin. 62 clones were randomly picked up and sequenced. As a result, 28 clones among them were found to be introduced mutations with threonine at position 203 being replaced by K, N, H, R or L and they produced non-fluorescent proteins. On the other hand, 34 clones produced fluorescent proteins, of which 23 contained mutants with threonine at position 203 was replaced by P, V, I, W, F, S, C or Y. In particular, a remarkable change in spectrum was observed in the clone in which threonine at position 203 was mutated into tyrosine. This mutation is referred to as ECGFP (enhanced cyan-green fluorescent protein) because its overall excitation and emission spectra were intermediate between those of ECFP and EGFP. The amino acid sequence of ECGFP is shown in SEQ ID NO.9.

The fluorescence excitation and emission spectra of ECGFP are shown together with those of ECFP and EGFP in Fig. 3. In the figure, the fluorescence excitation spectrum of ECGFP is shown by Curve A (broken line) and the emission spectrum of ECGFP by Curve B (solid line). Also, Fig. 3 shows the fluorescence excitation spectrum of EGFP by Curve C (broken line), the emission spectrum of EGFP by Curve D (solid line), the fluorescence excitation spectrum of ECFP by Curve E (broken line) and the emission spectrum of ECFP by Curve F (solid line).

As is seen from Fig. 3, the excitation and emission peaks of ECGFP (455 nm and 506 nm, respectively) were fairly close to those of the long wavelength component of ECFP (452 nm and 505 nm, respectively). This suggests that the $\pi \cdot \pi$ interaction favored the quantum state in which the longer wavelength predominated. Fig. 3 also indicates that ECGFP and ECFP have similar emission peaks (ECGFP, 506 nm and ECFP, 511 nm), but that ECGFP exhibits a larger Stokes' shift (the gap in wavelength between

excitation and emission peaks). In ECGFP, since it has a larger Stokes' shift than ECFP or EGFP, the fluorescent signal and excitation light can be easily separated in observation of the illuminated fluorescence.

Further, ECGFP, EGFP and ECFP were examined for pH resistance.

Firstly, using any of the following materials were used to prepare a series of pH buffers were prepared such that the pH values fell within the range of 3.8 to 11: 50 mM acetate (pH 3.8 to 5.6), $\text{Na}_2\text{HPO}_4\text{-Na}_2\text{H}_2\text{PO}_4$ (pH 5.5 to 7.0), 2-(N-morpho) ethane sulfonate (pH 5.5 to 6.0), 3-(N-morpho) propane sulfonate (pH 6.5 to 7.0), HEPES (pH 7.0 to 7.9), glycine (pH 9.0 to 10.0) and phosphate (pH 11). A solution for examining the pH resistance was prepared by concentrating protein in a weak buffer solution and then mixing said weak buffer solution with a buffer solution at a certain pH in an equivalent amount.

The absorbance and emission wavelength spectra measured using the resultant solutions are shown in Fig. 4 and Fig. 5. In Fig. 4, the representation is as follows: (A): the absorbance spectrum of EGFP, (B): the emission spectrum of EGFP, (C): the absorbance spectrum of ECFP, (D): the emission spectrum of ECFP, (E): the absorbance spectrum of ECGFP and (F): the emission spectrum of ECGFP. Further, in Fig. 5, (A) represents the re-plotted spectra of (A), (C) and (E) in Fig. 4, whereas (B) represents the re-plotted spectra of (B), (D) and (F) in Fig. 4.

Moreover, regarding EGFP, ECFP and ECGFP, the quantum yield (Φ) and molar absorbance coefficient (ϵ) were measured. The values of the quantum yield (Φ) and molar absorbance coefficient (ϵ) were measured in 50 mM HEPES (pH 7.5). The quantum yield (Φ) of ECGFP was determined by comparison with the standard (fluorescein ($\Phi=0.91$) and ECFP ($\Phi=0.40$)). In order to measure the protein concentration for calculating the molar absorbance coefficient, a Bradford assay kit (Bio Rad) was used, with BSA (bovine serum albumin) as the standard material. The

results thereof are shown in Table 1.

Table 1

Designation	mutations	$\lambda_{\text{abs}} (\epsilon)$	$\lambda_{\text{em}} (\Phi)$	pKa
EGFP	F64L, S65T	488 (55-57)	507-509 (0.60)	6
ECFP	F64L, S65T, Y66W, N146L, M153T, V163A, N164H	434 (33.9) 452	476 (0.40) 505	4-6
ECGFP	F64L, S65T, Y66W, N146L, M153T, V163A, N164H, T203Y	463 (23.9)	506 (0.14)	< 4

In the above Table 1, λ_{abs} represents the peak of the absorbance spectrum in units of nanometers. ϵ in parentheses is the molar absorbance coefficient in units of $10^3 \text{M}^{-1} \text{cm}^{-1}$. λ_{em} represents the peak of the emission spectrum in units of nanometers. Φ in parentheses is the quantum yield. Ka is the acid dissociation constant, obtained with $\text{pKa} = -\log_{10} \text{Ka}$. pKa was calculated from Fig. 5 (B).

Since the absolute brightness is calculated as the product of ϵ and Φ , Fig. 1 indicates that the absolute brightness of ECGFP is dimmer than those of EGFP and ECFP around neutral pH. It is known that almost all GFP mutants are quenched to some extent by acidic pH.

On the other hand, as clearly seen from Figs. 4 (A) and (B), EGFP showed an acidification-dependent decrease in the absorbance peak at 488 nm and a concomitant increase in the absorbance peak at 398 nm. In EGFP, the emission (511-nm peak) and excitation (488-nm peak) spectra decreased with decreasing pH, but the fluorescence excitation spectrum showed no compensating increase at 398 nm. This indicates that the species absorbing at 398 nm was non-fluorescent. As shown in Figs. 4 (A) and (B), EGFP exhibited a pKa of 6.0 for both its absorbance and emission spectra, indicating that as shown in Figs. 5 (A) and (B), its absorbance is pH-sensitive but its quantum yield is not.

